软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2014,25(Suppl.(1)):1−10

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

物联网位置隐私保护综述*

孙利民[1], 李 红[1,2], 王笑寒[3], 何云华[1,4]

[1](物联网信息安全技术北京市重点实验室(中国科学院 信息工程研究所),北京 100093)

[2](中国科学院大学,北京 100049)

[3](北京大学 软件与微电子学院,北京 102600)

[4](西安电子科技大学 计算机学院,陕西 西安 710071)

通讯作者: 孙利民, E-mail: sunlimin@iie.ac.cn

摘 要: 位置信息是物联网感知信息的基本要素之一,也是物联网提供基于位置服务的前提.位置信息在带来服务便利的同时,其泄露也带来诸多威胁.物联网位置隐私保护已成为当前的研究热点之一.综述了物联网位置隐私保护领域现有的工作,阐述了物联网位置隐私保护的目标与挑战,重点介绍物联网在定位过程、基于位置服务以及边信息中的位置隐私泄露方式及对应的位置隐私保护机制,并探讨了物联网位置隐私保护技术未来的发展方向.

关键词: 物联网;位置隐私保护;定位;基于位置服务;边信息

中文引用格式: 孙利民,李红,王笑寒,何云华.物联网位置隐私保护综述.软件学报,2014,25(Suppl.(1)):1−10. http://www.jos.org.cn/1000-9825/14001.htm

英文引用格式: Sun LM, Li H, Wang XH, He YH. Survey on the location privacy preservation in the Internet of things. Ruan Jian Xue Bao/Journal of Software, 2014,25(Suppl.(1)):1−10 (in Chinese). http://www.jos.org.cn/1000-9825/14001.htm

Survey on the Location Privacy Preservation in the Internet of Things

SUN Li-Min[1], LI Hong[1,2], WANG Xiao-Han[3], HE Yun-Hua[1,4]

[1](Beijing Key Laboratory of IOT Information Security (Institute of Information Engineering, The Chinese Academy of Sciences), Beijing 100093, China)

[2](University of Chinese Academy of Sciences, Beijing 100049, China)

[3](School of Software & Microelectronics, Peking University, Beijing 102600, China)

[4](School of Computer Science, Xidian University, Xi'an 710071, China)

Corresponding author: SUN Li-Min, E-mail: sunlimin@iie.ac.cn

Abstract: In the Internet of Things, location information is considered as not only one of the basic elements of sensing information collected by sensors but also an essential prerequisite for providing location based services. On the one hand, location information can facilitate people's lives, but on the other hand, such information exposure may cause great harm to the users. The location privacy preservation for Internet of Things has become one of the hottest topics in the academic world. This paper first states the objectives and challenges of achieving location privacy preservation for the Internet of Things. Then, it provides an overview on the location privacy leakages and the corresponding privacy preserving techniques in localization, location based services and side information. Finally, it discusses the future directions on location privacy preservation for the Internet of Things.

Key words: Internet of things; location privacy; localization; location based services; side information

物联网通过亿万物体对自身状态和环境的感知,实现人、机、物的响应与互动,为人们的生活和生产提供便捷、智慧的服务.位置信息是物联网感知信息的基本要素之一,也是物联网提供基于位置服务的前提.运输工

* 基金项目: 国家自然科学基金(61472418); 国家高技术研究发展计划(863)(2013AA014002)
收稿时间: 2014-05-10; 定稿时间: 2014-08-26

具的导航、敏感/贵重物资的运输跟踪、交通优化调度、叫车服务等诸多应用都会用到位置信息.在日常生活中,人们通过当前位置查找附近的加油站、餐馆、商场等.基于位置服务受到了国家的高度重视,我国科技部专门制定了《导航与位置服务科技发展"十二五"专项规划》.美国著名市场调研机构 Global Industry Analysts 预测[1],到 2015 年,基于位置服务的全球市场份额将超过 200 亿美元,用户量超过 10 亿.

位置信息在带来服务便利的同时,其信息泄漏也将带来诸多威胁.攻击者能够根据位置信息推断出用户的兴趣爱好、运动模式、健康状况等个人隐私信息,据报道,美国 Sense Network 公司[2]每天处理超过 40 亿条位置数据,能够提取用户生活习惯、年龄、收入等属性信息;位置信息的泄漏还可能导致用户被跟踪、遭到人生攻击等更为严重的后果,如 2013 年 1 月,深圳宝安职业技术学校高二女生因为在微博上泄漏自己的位置,而导致被跟踪杀害;位置隐私的泄露甚至会给国家安全造成威胁,据报道,2007 年美军飞行员使用了手机位置共享服务,导致 4 架直升机直接被基地组织摧毁.

1 物联网位置隐私保护的目标与挑战

隐私是指不愿告诉他人的或不愿公开的个人的事[3,4],即谁(Who)在什么时间(When)什么地点(Where)做什么事情(What),如图 1 所示.物联网位置隐私是一种特殊的隐私,指在物联网中用户不愿被外界所知晓的与位置相关的信息,如医院、酒吧等敏感位置,以及位置信息所揭露的个人信息,如家庭地址、健康状况等.根据用户的位置是否连续,物联网位置隐私分为单个位置的隐私和连续轨迹的隐私.单个位置的隐私是指谁(Who)到过什么地方(Where),而轨迹隐私则是指谁(Who)在某个时间段内(Time)的运动轨迹(Where).

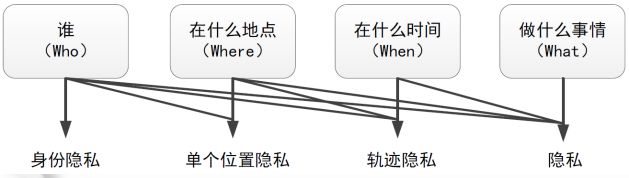


Fig.1 Concept of privacy in the Internet of things

图1 物联网位置隐私的概念

1.1 物联网位置隐私保护的目标

物联网位置隐私保护的目标是防止他人在用户不知情时获取用户在过去或现在的位置或轨迹信息,具体的目标可表示为:(1) 增加用户身份的不确定性,使攻击者不能确定用户的身份;(2) 增加位置的不确定性,使攻击者不能确定用户具体的位置;(3) 消除用户身份-位置之间的关联性,使攻击者不能将用户及其访问的位置关联起来.

设计物联网隐私保护机制时,还需兼顾数据的可用性以及系统的高效性.隐私保护机制通常需要在原有系统中加入复杂的算法,对用户的身份和位置数据进行处理,虽然提高了用户位置隐私的保护度,却降低了数据的可用性以及系统的高效性.因此,在设计隐私保护机制时,需在位置隐私保护与数据可用性、系统高效性之间确定一个平衡点.

1.2 物联网位置隐私保护的挑战

物联网中实现位置隐私保护存在一定的难度:(1) 物联网中存在很多泄露用户位置隐私的方式;(2) 位置隐私保护和基于位置服务是一对矛盾体,基于位置服务的服务质量越高,用户的位置隐私往往就越容易泄露;(3) 不同用户对位置隐私保护的要求不一样,同一用户在不同地点、不同环境下对位置隐私保护的要求也可能不一样;(4) 物联网中设备的能量、带宽等资源往往有限,需要轻量级的隐私保护机制.这些都对研究和设计位置隐私保护机制提出了挑战.

2 物联网位置隐私的泄露及防护机制

物联网中位置泄露的方式可分为 3 种:(1) 定位过程中的位置隐私泄露.物联网设备往往因自身资源有限,将定位过程外包给定位服务器,依靠定位服务器计算设备的位置.定位服务器在提供定位服务的同时,也获取了用户的位置隐私.(2) 在基于位置服务(location based service,简称 LBS)过程中的位置隐私泄露.在服务时,用户通常上传自己的位置信息,向 LBS 服务提供商请求与位置相关的服务,位置服务提供商因此获得用户的位置隐私.(3) 边信息(side information)中的位置隐私泄露.物联网产生大量与位置无关的数据(即边信息),攻击者可以从这些信息中推测出用户的位置隐私.

2.1 定位过程中的位置隐私泄露与防护

物联网中用户获取位置的方式分为以用户为中心的定位和以服务器为中心的定位两类.在以用户为中心的定位算法中,用户根据被动接收到的周围环境信号(如 GPS 信号等)估计自己当前的位置.以用户为中心的定位包括 GPS 定位和基于惯性传感器的定位等.在以用户为中心的定位中,用户的位置由自己计算,因此用户的位置隐私不容易泄露.在以服务器为中心的定位算法中,用户将接收到的周围环境信号发送给定位服务器,依靠服务器计算自己的位置.以服务器为中心的定位算法在用户端的开销较小,适合用于资源受限的物联网设备上,但定位服务器在提供定位服务的同时也获取了用户的隐私.

2.1.1 基于 WiFi 指纹定位的位置隐私及其防护机制

基于 WiFi 指纹的定位[5]是以服务器为中心的定位算法的典型代表之一.基于 WiFi 指纹的定位过程一般分为离线训练和在线定位两个阶段.在离线训练阶段,服务提供商在感兴趣的区域中选取若干参考点,测量这些参考点的 WiFi 指纹(周围 AP 的信号强度值),并将参考点的位置及其 WiFi 指纹存入 WiFi 指纹数据库.在线定位阶段,用户将当前位置的 WiFi 指纹发送给服务提供商,服务提供商通常找出与用户 WiFi 指纹欧式距离最近的 k 个指纹,并根据这 k 个指纹的位置估计用户的当前位置.这种基于 WiFi 指纹的定位往往存在位置隐私问题,定位服务提供商在为用户提供定位服务的同时,获取了用户位置.

Li 等人[6]提出了一种基于 Paillier Cryptosystem 加法同态性的位置隐私保护机制.用户给服务提供商发送加密后的 WiFi 指纹;服务提供商利用 Paillier Cryptosystem 的加法同态性,在密文空间上计算用户 WiFi 和指纹库中 WiFi 指纹的欧式距离;用户解密欧式距离,找出 k 个位置估计当前位置.具体过程如图 2 所示,基于同态加密的隐私保护机制将 WiFi 指纹定位分为预处理阶段、准备阶段、欧式距离计算阶段和位置估计阶段.预处理阶段,服务提供商从 WiFi 指纹数据库中抽取元数据,并发送给客户端.元数据中包含 WiFi 指纹采集地点、WiFi 指纹在数据库中的序号等信息.准备阶段,用户通过 Paillier Cryptosystem 生成一对公密钥,并将公钥及用公钥加密后的 WiFi 指纹发送给服务器.欧氏距离计算阶段,服务器在密文空间上计算用户指纹与 WiFi 指纹库中每个指纹之间的欧式距离,并将计算结果返回给用户.位置估计阶段,用户首先将接收到的欧式距离解密,找出 k 个最小的距离,然后根据 k 个最小的距离和元数据估算自己当前的位置.

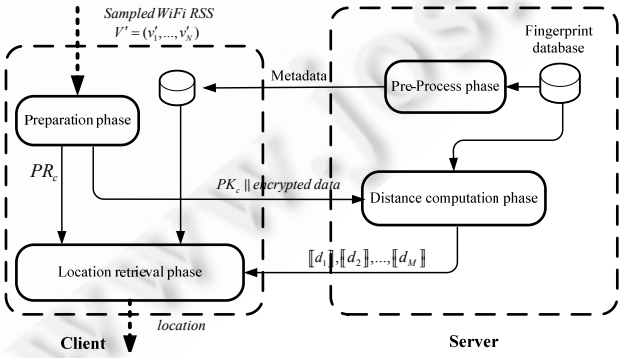


Fig.2 Workflow of the homomorphic encryption based privacy-preserving scheme

图2 基于同态加密的隐私保护机制的流程

2.1.2 基于测距的定位的位置隐私及其防护机制

基于测距的定位是物联网常用的定位算法之一.基于测距的定位算法通常分为两个步骤.首先,利用 TOA (time of arrival),TDOA(time difference of arrival)等方法测量用户位置 $x_0=(x_{01},...,x_{0n})$ 到第 i 个参考节点(anchor node)A_i 的距离为 d_{0i},其中 A_i 的位置为 $x_i=(x_{i1},y_{in})(i=1\sim m)$; 然后通过最小二乘法估计用户的位置 $\hat{x}_0=(A^TA)^{-1}A^Tb$,其中,

$$A=2\begin{bmatrix} x_{m1}-x_{11} & x_{mn}-x_{1n} \\ x_{m1}-x_{21} & x_{mn}-x_{2n} \\ \cdots & \cdots \\ x_{m1}-x_{m-1,1} & x_{mn}-x_{m-1,n} \end{bmatrix},\ b=\begin{bmatrix} \sum_{j=1}^{n}(x_{mj}^2-x_{1,j}^2)-(d_{0m}^2-d_{01}^2) \\ \sum_{j=1}^{n}(x_{mj}^2-x_{2,j}^2)-(d_{0m}^2-d_{02}^2) \\ \\ \sum_{j=1}^{n}(x_{mj}^2-x_{m-1,j}^2)-(d_{0m}^2-d_{0,m-1}^2) \end{bmatrix}.$$

在基于测距的定位算法中,用户的位置通常由定位服务器计算.如果位置信息发生泄露,用户的隐私将受到威胁.另外,定位服务器需要知道每个参考节点的位置才能计算用户的位置.如果参考节点的位置信息遭到泄露,攻击者可以对参考节点发动位置欺骗等攻击[7,8].

Shu 等人[9]为解决基于测距的定位算法的隐私问题,提出了一种基于安全最小二乘估计的隐私保护机制.在该机制中,每个参考节点不直接透露自己的位置,仅进行部分计算,用户汇总每个参考节点的计算结果,并最终估计自己的当前位置.由于用户不能通过中间计算结果推测出参考节点的具体位置,且最后由自己计算当前的位置,因此该机制能够同时保护参考节点和用户的位置隐私.

2.1.3 A-GPS 定位的位置隐私泄露

在传统的 GPS 定位机制中,移动终端需首先搜索当前区域内可用的卫星,然后测量到每个可用卫星之间的伪距,并根据伪距估计当前的位置.GPS 定位属于以用户为中心的定位技术,不存在泄漏用户位置隐私的风险,但首次搜索卫星的时间较长,定位精度容易受到建筑物、树木等障碍物的影响.

A-GPS(assisted-GPS)[10]是将 GPS 定位和无线蜂窝网定位相结合的定位技术.A-GPS 的基本原理是通过一个后端的定位服务器收集卫星的星历信息;当终端发出定位请求时,后端定位服务器根据终端使用的蜂窝网基站的 CELL ID 判断终端的大致区域,然后将该区域可见卫星的相关信息发送给终端;终端接收可见卫星的广播信号,计算与每个可见卫星间的伪距,并将伪距信息通过蜂窝网发送给后端的定位服务器;定位服务器根据伪距信息计算终端的位置坐标,并将结果返回给终端.A-GPS 首次搜索卫星速度较快,定位精度较高,但定位服务器在给用户提供定位服务时获取了用户的位置隐私.人们通常采用构造虚假卫星伪距的方式保护用户的位置隐私.

2.2 LBS中的位置隐私泄露与防护

基于位置的服务是物联网位置隐私泄露的主要方式之一.当用户请求基于位置的服务时,如查找最近的加油站,通常需要将自己当前的位置或移动轨迹发送给服务提供商,服务提供商因此获取了用户的位置隐私.服务提供商还可能将用户轨迹信息发布,用于科研、城市规划等,攻击者可通过数据挖掘等技术推测用户的隐私.根据用户的位置是否连续,LBS 中的位置隐私保护机制可分为单点位置的隐私保护和移动轨迹的隐私保护两类.

2.2.1 单点位置隐私保护机制

为了防止用户在获取 LBS 服务时位置隐私泄露,研究人员对单点位置隐私保护技术展开了研究.位置隐私防护的目标是:身份保护,服务器提供商能获取位置但不知是谁请求服务;位置保护,服务器知道谁在请求服务但不能获取其准确位置;身份及位置防护,服务提供商不能获知谁在哪里请求服务.按照位置隐私保护目标,现有技术大致可分为 3 类.

(1) k-匿名混淆(k-anonymity cloaking).k-匿名是隐私保护最常用的技术之一,其概念最初用于关系数据库中数据发布隐私保护.Gruteser 等人[11]将 k-匿名的概念应用到位置隐私保护上,基本思想是使某个区域至少有 k 个用户,这 k 个用户之间不能直接通过 ID 来区别.k-匿名的实现方式可分为集中式和分布式两种.集中式方法在移动用户与 LBS 服务提供商之间加入可信第三方,由它完成匿名的处理和查询工作,如图 3 所示.基于四叉树方

法[11],由第三方自顶而下划分整个空间,直到查询用户所在区域的用户数小于 k,将上一层划分区域作为匿名区域.然而,实际中很难找到用户信任的第三方,而且攻击者一旦成功俘获第三方,所有用户的隐私将遭受侵害.

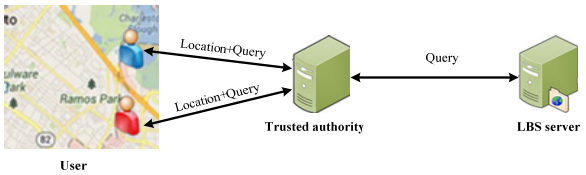


Fig.3 Centralized approach

图3 集中式方法

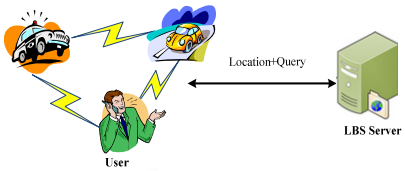


Fig.4 Distributed approach

图4 分布式方法

与集中式方法不同,分布式方法不依赖于可信第三方,仅通过用户相互之间的协作构成 k-匿名集,如图 4 所示.Chow 等人[12]提出分布式 P2P 的 k 匿名混淆算法,用户请求服务之前,先通过单跳或多跳通信与周围 k–1 个用户形成组,将自己的位置混淆到包含组中的所有用户的匿名区域中,然后通过挑选的组代理来完成请求服务.当某个区域中的用户数少于 k 时,Lu 等人[13]提出通过构造虚假用户来满足用户数的要求,然而构造和上传虚假的 LBS 请求会带来额外开销,用户期望其余用户完成.Liu 等人[14]通过贝叶斯博弈分析了用户构造虚假用户、上传虚假 LBS 请求的行为,提出了基于博弈理论的虚假用户构造机制.Yang 等人[15]提出 k-匿名激励机制,由关心隐私用户给不关心隐私用户给予相应报酬,来激励其他用户构造假用户,以满足 k-匿名要求.

(2) 位置模糊.该技术使用圆形、矩形等区域替代用户的真实位置[16],或向用户的真实位置中添加可控的噪声,使得攻击者很难推断出用户真实的位置[17,18].Xu 等人[16]提出基于用户 feeling 的模型来表达用户隐私需求,根据上报用户感觉合适的区域来划定区域,如用户不介意将商场作为公开区域,但介意将办公楼作为公开区域.

另一类假位置方法,通过向真实位置添加可控的噪声得到假位置.由于隐私保护度依赖于用户密度,用户多应加入较少噪声,用户少应引入较大噪声.Pingley 等人[17]提出了基于 Hibert 曲线的位置扰乱方法,将不同密度用户映射到等密度的 Hibert 曲线上,然后加入统一的噪声,即密度小时噪声多,密度大时噪声少.Andrés 等人[18]提出基于地理位置差分隐私扰乱方法,通过添加平面 Laplace 噪声,保证攻击者从观察到的位置信息得不到关于用户位置的任何信息,为用户提供强隐私保护.

(3) 密码学方法.该方法将用户的位置、兴趣点加密后,在密文空间计算和搜寻查询结果,服务提供商不能获取查询内容和查询结果的具体内容.安全多方计算是常用的隐私保护算法,在 n 个互相不信任的参与方中进行协同计算,而不侵犯任何一方的隐私信息,可通过同态加密、双线性对等实现.然而,LBS 服务提供商和移动用户的加解密操作以及密文空间上的匹配操作,都带来较大的计算开销.Shao 等人[19]采用代理重加密技术将服务器和用户端的加密和匹配操作放在云端进行,云服务器只需验证位置是否匹配,而对 LBS 提供商和用户的真实位置并不知晓.

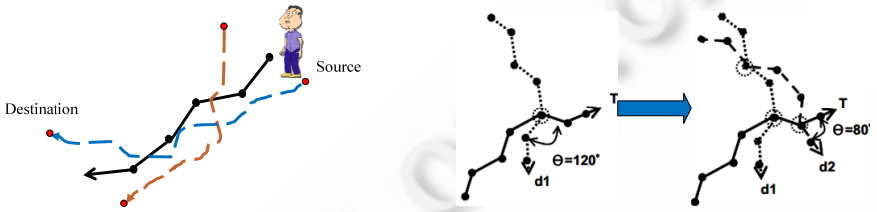
2.2.2 连续轨迹的隐私保护机制

在导航、寻找加油站等连续的 LBS 服务中,用户需连续上传位置到 LBS 服务器,移动用户的隐私可能通过实时运行轨迹而暴露;轨迹发布用于,如城市道路规划、移动网络性能优化等应用,攻击者通过数据挖掘分析,可以推断出个人的兴趣爱好、行为模式、社会习惯等隐私信息.MIT 的研究人员[20]通过对 150 万人长达 15 个月的移动轨迹的研究发现,用户的移动轨迹具有高度的唯一性,通过 4 个位置点能够成功识别 95%用户的轨迹.

单点位置的匿名和模糊不能保证节点移动轨迹的隐私,攻击者可能将用户时间顺序上的多个位置或区域信息连接起来,从而得到用户在某一段时间内的运动轨迹.轨迹上传或发布之前,通常用随机 ID 替换用户真实标识.然而,Ma 等人[21]指出,这种匿名方式不足以保护用户隐私,他们提出的主动和被动攻击,能以较大概率识别出用户的轨迹.Srivatsa 等人[22]利用社交网络关系图高效识别了 80%的用户轨迹.随着研究的不断深入,大量的轨迹隐私保护方法被提了出来,这些隐私保护技术的目标可概括为:身份保护,防止攻击者获取轨迹的真实拥有者;时空相关性防护,防止攻击者根据轨迹本身的时空相关性跟踪用户.现有轨迹隐私保护技术大致可分为以下 3 类[3].

(1) 假数据.该方法通过添加假轨迹对原始数据进行干扰,产生逼真、混淆性的假轨迹数据,增加攻击者识别的难度.在连续LBS场景中,可信第三方产生或用户自己产生虚假轨迹,与用户真实轨迹一起向LBS服务器请求服务.Xu 等人[23]提出由可信第三方根据用户历史轨迹,建立 footprint 数据库,每次请求服务,采用数据库中的轨迹数据产生假轨迹.Pingley 等人[24]考虑由用户自己产生假查询,基于位置服务中查询的合理性、节点的移动模式,来产生更逼真的假位置查询,然而这将产生大量的通信开销和存储开销.

在轨迹发布场景中,轨迹数据发布之前,由 LBS 服务器产生一些假轨迹,使得攻击者看到的轨迹数据增多,加大了识别真实轨迹的难度.You 等人[25]提出两种假轨迹产生方法:随机法,由 LBS 服务器随机生成一条连接起点到终点、连续运行且运行模式一致的假轨迹,如图 5(a)所示;旋转法,以移动用户的真实轨迹为基础,以真实轨迹中的某些采样点为轴点进行旋转,旋转后的轨迹为生成的假轨迹,如图 5(b)所示.旋转点的选择和旋转角度的确定需要和信息扭曲度进行关联权衡.



(a) Random pattern scheme (b) Rotation pattern scheme

Fig.5 Schemes based on dummy trajectories

图5 假轨迹产生方法

(2) 抑制法.该方法抑制轨迹中某些位置,如敏感位置或易泄露用户隐私的位置,增加攻击者识别用户轨迹的难度,仅适用于了解攻击者拥有某种特定背景知识的情形,隐私保护能力有限.在连续 LBS 场景中,用户通过限制上传地点,防止攻击者进行跟踪攻击.Mix 域方法让用户进入 Mix 域时更新假名,不发送位置信息到服务器,使得攻击者难以区分 Mix 域中的用户(如图 6 所示),理想的 Mix 域要求 k 个用户同时在 Mix 域中,这些用户进出 Mix 的时间是随机的,进出 Mix 域的转移概率是均匀的.Palanisamy 等人[26]考虑车辆密度、Mix 形状、位置粒度、移动限制等因素,建立适应于道路网的 Mix 域模型.Liu 等人[27]提出了最优放置 Mix 域的方法.Hoh 等人[28]提出在地图中设定的标记点,节点在这些标记点上传位置信息,标记点的放置确保节点隐私保护的最小需求距离,并且避免暴露一些隐私敏感的位置.

在轨迹发布场景中,可信服务器抑制轨迹中某些点之后再发布,并保证发布数据的可用性.Teriovitis 等人[29]指出应抑制轨迹中的敏感点,如诊所、酒吧等,以及轨迹中会泄露其他信息的位置点,如可唯一识别用户轨迹的位置点和可区分于其他轨迹的位置点.Mohammed 等人[30]提出一种$(K,C)_L$ 隐私模型来指导选择抑制点,保证轨迹集中的每个长度为 L 的子序列至少与其他 K–1 个子序列的数据记录不可区分,同时每个匿名组中的敏感位置比率不超过概率 C.

(3) 泛化法.该方法将轨迹上的位置点都泛化为对应的匿名区域,在信息扭曲度最小的情况下达到轨迹 k-匿名,该类方法数据失真相对较小,但实现优化轨迹匿名的开销大.在连续 LBS 场景中,位置匿名服务器将用户连续的位置点泛化为对应匿名区域后,发给 LBS 服务提供商请求服务.Chow 等人[31]提出基于组的空间混淆技术,将用户与附近 k–1 个邻居形成组,在接下来的一段时间内,通过上传组区域来获取服务.Pan 等人[32]提出根据节点的速度、方向,选择相似移动特性的节点形成组.

针对轨迹发布场景,Abul 等人[33]提出(k,δ)匿名泛化方法,对同一段时间内的轨迹数据,按欧式距离进行贪心聚类,直到每个类中至少含有 k 条轨迹,然后把不在匿名区域(半径为δ)的点在最小信息扭曲的前提下,移动到匿名区域内,如图 7 所示.Nergiz 等人[34]聚类匿名后,在匿名区域内随机选择点进行重构轨迹,发布重构轨迹,具有原始轨迹特性,数据失真相对较小.

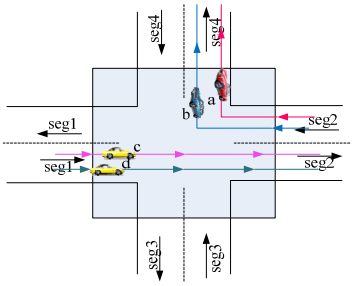


Fig.6 Mix zone
图6 Mix 域

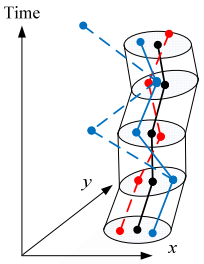


Fig.7 (k,δ)-Anonymity
图7 (k,δ)匿名泛化方法

2.3 边信息中的位置隐私泄露与防护

除了定位过程和基于位置服务中存在位置泄露外,与位置无直接关联的数据也可能泄露位置相关信息.当用户访问某些地点后,会在某些设备上留下特定的信息,或致使设备的测量数据发生特定的变化.这些与位置无直接关联却能从中推测出用户位置隐私的数据被称为边信息.物联网的感知层存在大量的感知节点,这些节点能够实时、准确地感知用户在物理空间的活动与状态,因此物联网存在大量、丰富的边信息.攻击者可以通过数据挖掘、大数据分析等手段对这些边信息进行分析,获取用户的位置隐私.下面将介绍物联网中一些典型的边信息及其防护机制.

2.3.1 移动设备的 WiFi 连接记录

智能手机、平板电脑等移动设备为了快速发现/切换 WiFi AP,或者发现隐藏的 WiFi AP,通常会保存 AP 的连接记录.当移动设备需要再次连接 AP 时,主动广播包含已连接过 AP 的 SSID 的探测请求(probe request),询问之前连接过的 AP 是否存在.攻击者通过监听移动设备的 AP 主动发现过程获取用户的 AP 连接记录;然后通过查询 Google、Skyhook 等公司采集的 WiFi AP 数据库,即可得到连接记录中 AP 的详细位置,进而推断出用户曾经到过的位置或轨迹.

Lindqvist 等人[35]提出了一种基于挑战-应答(challenge-response)的 AP 发现机制,防止攻击者通过窃听的方式获取用户的 WiFi 连接记录.在该机制中,移动设备和已连接过的 AP 之间共享一个密钥.当设备进行 AP 主动发现时,设备通过挑战应答协议发现之前连接过的 AP:设备产生一个随机数,作为“提问”发送给 AP;AP 将 SSID 和随机数合并,并用共享密钥加密签名作为“应答”返回给移动设备;移动设备比较计算结果和“应答”,如果两者相同,则证明用户之前连接过该 AP.Cox[36]提出,在移动设备与 AP 首次连接时协商一个密钥,移动设备在广播探测请求时用该密钥对 AP 的 SSID 进行加密.

2.3.2 设备通信过程中泄露的信号强度

由于无线信道的广播特性,攻击者通过窃听用户与 AP 之间的无线通信,获取用户设备的 MAC 地址并测量用户发送信号的信号强度.由于 MAC 地址具有唯一性,因此攻击者可以据此确定用户的身份.当同时在 4 个不同地方测量用户发送信号的功率时,攻击者通过三角定位法可估计出用户的具体位置.

为了防止攻击者通过信号强度值对用户进行定位,Jiang[37]提出了通过假名和减小设备发射功率来保护用户位置隐私的机制.移动设备与 AP 通信时不使用真实的 MAC 地址,而是采用由 AP 分配的假名.Wang[38]提出用户进入区域时,首先通过 Log-Normal 信号衰减模型估算周围 AP 的位置,然后利用智能天线调整在不同方向的发射功率,尽量使能监听到移动设备 WiFi 通信的 AP 数少于 4;或者调整移动设备在不同方向上的发射功率,使攻击者通过三角定位的误差达到最大.

2.3.3 智能电网中电表的数据

在智能电网中,智能电表周期性地记录用户的用电量,并将数据发送给电网公司.攻击者根据电表读数变化推测出用户是否在家,甚至推测出用户家中使用的电器、行为习惯等隐私信息.

Kalogridis[39]提出的 BE(best effort)算法利用蓄电池尽可能地使家庭用电量保持一个常数,家庭用电量少时

给蓄电池充电,用电量大时蓄电池放电.McLaughlin[40]把蓄电池的电量分为 ST(stable state),HR(high recovery)和 LR(low recovery)3 个状态.在每个状态下,电网供电量保持一个常数不变,攻击者观察到的电量变化呈现阶跃函数的形式,不能从用电数据的变化推测出用户的行为.Chow[41]提出了一种基于同态加密的隐私保护方案,用户对实时用电数据进行加密,电网公司根据密文计算用户的电费.

3 总结与展望

随着物联网技术的不断发展,物联网位置隐私问题逐渐受到人们的重视,已经成为阻碍物联网进一步推广的决定性因素之一.本文综述了物联网位置隐私保护领域现有的工作,总结了物联网位置隐私保护的目标与挑战,重点介绍了在定位过程、基于位置服务以及边信息中的位置隐私泄露方式及相应的位置隐私保护机制.

虽然物联网位置隐私保护在近几年得到了很多研究,但它仍是当前的热点之一,还存在许多问题有待进一步研究:(1) 基于位置服务中的位置隐私保护研究较为成熟,研究人员提出了很多通用的隐私保护机制,但总体上对应用场景的实际情况考虑较少,实用性较差.需结合特定 LBS 应用的特点,研究可靠的、实用的隐私保护机制.(2) 定位过程中的隐私保护是物联网位置隐私保护新的研究点.物联网的很多定位算法存在位置隐私泄露问题,亟需研究相应的位置隐私保护机制.(3) 物联网存在大量、丰富的边信息,这些边信息的隐私泄露构成位置隐私新的威胁(如手机加速度传感器的读数[42]).如何发现和预防新的边信息位置隐私泄露是物联网位置隐私保护研究的新挑战.(4) 现有的研究中缺乏严格的攻击模型,对攻击者所具有的知识没有做出定量化的定义.差分隐私[43]是数据库隐私领域提出的新模型,它定义了一个极为严格的攻击模型,在降低隐私泄露风险的同时保证了数据的可用性.如何应用差分隐私保护用户的位置隐私是物联网位置隐私保护新的研究点.(5) 通过大数据等相关技术,攻击者可以从多种渠道获得用户位置数据,并通过数据挖掘等手段推测用户的隐私.如何在大数据时代保护用户的位置隐私是物联网位置隐私保护研究所面临的新的挑战.

References:

[1] Global Industry Analysts. Location-Based Service (Lbs)–A Global Market Report. http://www.prweb.com/releases/location_based_services/LBS/prweb4370484.htm

[2] Sense Network. Sense Target: How it Works. https://www.sensenetworks.com/how-it-works/

[3] Huo Z, Meng XF. A survey of trajectory privacy-preserving techniques. Chinese Journal of Computers, 2011,34(10):1820−1830 (in Chinese with English abstract).

[4] Zhou SG, Li F, Tao YF, Xiao XK. Privacy preservation in database applications: A survey. Chinese Journal of Computers, 2009, 32(5):847−861 (in Chinese with English abstract).

[5] Bahl P, Padmanabhan VN. Radar: An in-building RF-based user location and tracking system. In: Proc. of the 19th Annual IEEE Int'l Conf. on Computer Communications (Infocom 2000). IEEE Computer Society Press, 2000. 775−784.

[6] Li H, Sun LM, Zhu HJ, Lu X, Cheng XZ. Achieving privacy preservation in WiFi fingerprint-based localization. In: Proc. of the 33rd Annual IEEE Int'l Conf. on Computer Communications (Infocom 2014). IEEE Computer Society Press, 2014. 2337−2345.

[7] Wang T, Yang Y. Analysis on perfect location spoofing attacks using beamforming. In: Proc. of the 32nd Annual IEEE Int'l Conf. on Computer Communications (Infocom 2013). IEEE Computer Society Press, 2013. 2778−2786.

[8] Yang J, Chen Y. Towards attack resistant localization under infrastructure attacks. Security and Communication Networks, 2012, 5(4):384−403.

[9] Shu T, Chen YY, Yang J, Williams A. Multi-Lateral privacy-preserving localization in pervasive environments. In: Proc. of the 33rd Annual IEEE Int'l Conf. on Computer Communications (Infocom 2014). IEEE Computer Society Press, 2014. 2319−2327.

[10] LaMance J, DeSalas J, Järvinen J. Assisted GPS: A low-infrastructure approach. GPS World, 2002,13(3):46−51.

[11] Gruteser M, Grunwald D. Anonymous usage of location-based services through spatial and temporal cloaking. In: Proc. of the 1st Int'l Conf. on Mobile Systems, Applications and Services (Mobisys 2003). ACM Press, 2003. 31−42.

[12] Chow CY, Mokbel MF, Liu X. A peer-to-peer spatial cloaking algorithm for anonymous location-based services. In: Proc. of the 14th Annual ACM Int'l Symp. on Advances in Geographic Information Systems (GIS 2006). ACM Press, 2006. 171−178.

[13] Lu H, Jensen CS, Yiu ML. Pad: Privacy-Area aware, dummy-based location privacy in mobile services. In: Proc. of the 7th ACM Int'l Workshop on Data Engineering for Wireless and Mobile Access (MobiDE 2008). ACM Press, 2008. 16−23.

[14] Liu XX, Liu KK, Guo LK, Li XL, Fang YG. A game-theoretic approach for achieving *k*-anonymity in location based services. In: Proc. of the 32nd Annual IEEE Int'l Conf. on Computer Communications (Infocom 2013). IEEE Computer Society Press, 2013. 2985−2993.

[15] Yang DJ, Fang X, Xue GL. Truthful incentive mechanisms for *k*-anonymity location privacy. In: Proc. of the 32nd Annual IEEE Int'l Conf. on Computer Communications (Infocom 2013). IEEE Computer Society Press, 2013. 2994−3002.

[16] Xu T, Cai Y. Feeling-Based location privacy protection for location-based services. In: Proc. of the 16th ACM Conf. on Computer and Communications Security (CCS 2009). ACM Press, 2009. 348−357.

[17] Pingley A, Yu W, Zhang N, Fu X, Zhao W. Cap: A context-aware privacy protection system for location-based services. In: Proc. of the IEEE Int'l Conf. on Distributed Computing Systems (ICDCS 2009). IEEE Press, 2009. 49−57.

[18] Andrés ME, Bordenabe NE, Chatzikoklakis K, Palamidessi C. Geo-Indistinguishability: Differential privacy for location-based systems. In: Proc. of the ACM Conf. on Computer & Communications Security (CCS 2013). ACM Press, 2013. 901−914.

[19] Shao J, Lu RX, Lin XD. FINE: A fine-grained privacy-preserving location-based service framework for mobile devices. In: Proc. of the 33rd Annual IEEE Int'l Conf. on Computer Communications (Infocom 2014). IEEE Computer Society Press, 2014.

[20] Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: The privacy bounds of human mobility. Nature, 2013,3(1):1−5.

[21] Ma CY, Yau DK, Yip NK, Rao NS. Privacy vulnerability of published anonymous mobility traces. IEEE Trans. on Networking, 2012,21(3):720−733.

[22] Srivatsa M, Hicks M. Deanonymizing mobility traces: Using social networks as a side-channel. In: Proc. of the ACM Conf. on Computer and Communications Security (CCS 2012). ACM Press, 2012. 628−637.

[23] Xu T, Cai Y. Exploring historical location data for anonymity preservation in location-based services. In: Proc. of the 27th Annual IEEE Int'l Conf. on Computer Communications (Infocom 2008). IEEE Computer Society Press, 2008. 1220−1228.

[24] Pingley A, Zhang N, Fu XW. Protection of query privacy for continuous location based services. In: Proc. of the 30th Annual IEEE Int'l Conf. on Computer Communications (Infocom 2011). IEEE Computer Society Press, 2011. 1710−1718.

[25] You TH, Peng WC, Lee WC. Protecting moving trajectories with dummies. In: Proc. of the IEEE Int'l Conf. on Mobile Data Management (MDM 2007). IEEE Press, 2007. 278−282.

[26] Palanisamy B, Liu N. MobiMix: Protecting location privacy with mix-zones over road networks. In: Proc. of the Int'l Conf. on Data Engineering (ICDE 2011). IEEE Press, 2011. 494−505.

[27] Liu XX, Zhao H, Pan M, Yue H, Li XL, Fang YG. Traffic-Aware multiple mix zone placement for protecting location privacy, In: Proc. of the 31st Annual IEEE Int'l Conf. on Computer Communications (Infocom 2012). IEEE Computer Society Press, 2012. 972−980.

[28] Hoh B, Gruteser M. Virtual trip lines for distributed privacy-preserving traffic monitoring. In: Proc. of the 6th Int'l Conf. on Mobile Systems, Applications, and Services (MobiSys 2008). ACM Press, 2008. 15−28.

[29] Terrovitis M, Mamoulis N. Privacy preservation in the publication of trajectories. In: Proc. of the IEEE Int'l Conf. on Mobile Data Management (MDM 2008). IEEE Press, 2008. 65−72

[30] Mohammed N, Fung BC, Debbabi M. Walking in the crowd: Anonymizing trajectory data for pattern analysis. In: Proc. of the Int'l Conf. on Information and Knowledge Management (CIKM 2009). ACM Press, 2009. 1441−1444.

[31] Chow CY, Mokbel MF. Enabling private continuous queries for revealed user locations. In: Proc. of the Int'l Symp. on Spatial and Temporal Databases (SSTD 2007). Springer-Verlag, 2007. 258−275.

[32] Pan X, Meng X, Xu J. Distortion-Based anonymity for continuous queries in location-based mobile services. In: Proc. of the Int'l Conf. on Advances in Geographic Information Systems (GIS 2009). ACM Press, 2009. 256−265.

[33] Abul O, Bonchi F, Nanni M. Never walk alone: Uncertainty for anonymity in moving objects databases. In: Proc. of the IEEE Int'l Conf. on Data Engineering (ICDE 2008). IEEE Press, 2008. 376−385.

[34] Nergiz M, Atzori M, Saygin Y, Guc B. Towards trajectory anonymization: A generalization-based approach. Journal Trans. on Data Privacy, 2009,2(1):47−75.

[35] Lindqvist J, Aura T, Danezis G, Koponen T, Myllyniemi A, Maki J. Privacy-Preserving 802.11 access-point discovery. In: Proc. of the ACM Conf. on Wireless Network Security (WiSec 2009). ACM Press, 2009. 123−130.

[36] Cox LP, Dalton A, Marupadi V. Smokescreen: Flexible privacy controls for presence-sharing. In: Proc. of the 5th Int'l Conf. on Mobile Systems, Applications and Services (MobiSys 2007). ACM Press, 2007. 233−245.

[37] Jiang T, Wang HJ, Hu YC. Preserving location privacy in wireless LANs. In: Proc. of the 5th Int'l Conf. on Mobile Systems, Applications and Services (MobiSys 2007). ACM Press, 2007. 246−257.

[38] Wang T, Yang YL. Location privacy protection from rss localization system using antenna pattern synthesis. In: Proc. of the IEEE Int'l Conf. on Computer Communications (INFOCOM 2011). IEEE Press, 2011. 2408−2416.

[39] Kalogridis G, Efthymiou C, Denic S, Lewis T, Cepeda R. Privacy for smart meters: Towards undetectable appliance load signatures. In: Proc. of the 1s Int'l Conf. on Smart Grid Communications (SmartGridComm 2010). IEEE Press, 2010. 232−237.

[40] McLaughlin S, McDaniel P, Aiello W. Protecting consumer privacy from electric load monitoring. In: Proc. of the 18th ACM Conf. on Computer and Communications Security (CCS 2011). ACM Press, 2011. 87−98.

[41] Shi E, Chan TH, Rieffel EG, Chow R, Song D. Privacy-Preserving aggregation of time-series data. In: Proc. of the 18th Annual Network & Distributed System Security (NDSS 2011). Internet Society Press, 2011.

[42] SFGate. Stanford researchers discover 'alarming' method for phone tracking. http://blog.sfgate.com

[43] Dwork C. Differential privacy: A survey of results. In: Proc. of the 5th Int'l Conf. on Theory and Applications of Models of Computation (TAMC 2008). Springer-Verlag, 2008. 1−19.

附中文参考文献:

[3] 霍峥,孟小峰.轨迹隐私保护技术研究.计算机学报,2011,34(10):1820−1830.

[4] 周水庚,李丰,陶宇飞,肖小奎.面向数据库应用的隐私保护研究综述.计算机学报,2009,32(5):847−861

孙利民(1966－),男,河南周口人,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为无线传感网,物联网及其安全.
E-mail: sunlimin@iie.ac.cn

李红(1989－),男,博士生,CCF 学生会员,主要研究领域为物联网位置隐私保护.
E-mail: lihong@iie.ac.cn

王笑寒(1987－),女,硕士生,主要研究领域为通信工程.
E-mail: pagewang@foxmail.com

何云华(1987－),男,博士生,CCF 学生会员,主要研究领域为车载自组织网络的安全,隐私保护.
E-mail: heyunhua610@163.com